AMENDMENTS TO THE CLAIMS

This listing of claims replaces all prior versions and listings of claims in the application:

Listing of claims:

1.   (Currently Amended) A method comprising:

receiving input from a user identifying at least two portions of a first set of audio signals as being of interest to the user;

processing, by a query recognizer of a word spotting system, each identified portion of the first set of audio signals to generate a corresponding subword unit representation of the identified portion;

forming, by ~~a~~ the query recognizer of ~~a~~ the word spotting system, a representation of a spoken event of interest, wherein the forming includes ~~receiving an indication that a spoken event in a first set of audio signals is of interest to a user, identifying two or more instances of the spoken event of interest in a the first set of audio signals, and representing each identified instance of the spoken event of interest in the representation of the spoken event of interest using at least one sequence of subword units~~ combining the subword unit representations of the respective identified portions of the first set of audio signals;

accepting, by a word spotting engine of the word spotting system, data representing ~~the~~ unknown speech in a second audio signal; and

locating, by the word spotting engine of the word spotting system, putative instances of the spoken event of interest in the second audio signal using the representation of the spoken event of interest, wherein the locating includes identifying time locations of the second audio signal at which the spoken event of interest is likely to have occurred based on a comparison of the data representing the unknown speech with the representation of the spoken event of interest.

2.   (Currently Amended) The method of claim 1 wherein <u>processing each identified</u> <u>portion of the first set of audio signals</u> ~~forming the representation of the spoken event of~~ ~~interest~~ comprises applying a computer-implemented speech recognition algorithm to data representing the first set of audio signals.

3.   (Original) The method of claim 1 wherein the subword units include linguistic units.

4.   (Previously Presented) The method of claim 2 wherein locating the putative instances includes applying a computer-implemented word spotting algorithm configured using the representation of the spoken event of interest.

5.   (Previously Presented) The method of claim 4 further comprising selecting processing parameter values of the speech recognition algorithm for application to the data representing the first set of audio signals according to characteristics of the word spotting algorithm.

6.   (Previously Presented) The method of claim 5 wherein the selecting of the processing parameter values of the speech recognition algorithm includes optimizing said parameters according to an accuracy of the word spotting algorithm.

7.   (Previously Presented) The method of claim 5 wherein the selecting of the processing parameter values of the speech recognition algorithm includes selecting values for parameters including one or more of an insertion factor, a recognition search beam width, a recognition grammar factor, and a number of recognition hypotheses.

8.   (Previously Presented) The method of claim 1 wherein the representation of the spoken event of interest defines a network of subword units.

9.   (Previously Presented) The method of claim 8 wherein the network of subword units is formed by multiple sequences of subword units that correspond to different paths through the network.

10.   (Previously Presented) The method of claim 1 wherein forming the representation of the spoken event of interest includes determining an n-best list of recognition results.

11.   (Previously Presented) The method of claim 10 wherein each sequence of subword units in the representation corresponds to a different one in the n-best list of recognition results.

12.   (Previously Presented) The method of claim 1, further comprising accepting first audio data representing utterances of the event of interest spoken by a user, and processing the first audio data to form a processed query.

13.   (Currently Amended) The method of claim_1, further comprising accepting a selection by the user of portions of stored data from the first set of audio signals, and processing the portions of the stored data to form a processed query.

14.   (Previously Presented) The method of claim 13 further comprising, prior to accepting the selection by the user, processing the first set of audio signals according to a first computer-implemented speech recognition algorithm to produce the stored data.

15.   (Previously Presented) The method of claim 14 wherein the first speech recognition algorithm produces data related to presence of the subword units at different times in the first set of audio signals.

16.   (Previously Presented) The method of claim 14, further comprising applying a second speech recognition algorithm to the processed query.

17. (Currently Amended) A computer-readable medium storing instructions for causing a processing system to:

> receive input from a user identifying at least two portions of a first set of audio signals as being of interest to the user;

> process each identified portion of the first set of audio signals to generate a corresponding subword unit representation of the identified portion;

> form a representation of a spoken event of interest, wherein the instructions for causing the processing system to form the representation include instructions for ~~receiving an indication that a spoken event in a first set of audio signals is of interest to a user, identifying two or more instances of the spoken event of interest in the first set of audio signals, and representing each identified instance of the spoken event of interest in the representation of the spoken event of interest using at least one sequence of subword units~~ combining the subword unit representations of the respective identified portions of the first set of audio signals;

> accept data representing ~~the~~ unknown speech in a second audio signal; and

> locate putative instances of the spoken event of interest in the second audio signal using the representation of the spoken event of interest, wherein the instructions for causing the processing system to locate the putative instances include instructions for identifying time locations of the second audio signal at which the spoken event of interest is likely to have occurred based on a comparison of the data representing the unknown speech with the specification of the spoken event of interest.

18.   (Currently Amended) A system comprising:

a speech recognizer for <u>receiving input from a user identifying at least two portions of a first set of audio signals as being of interest to the user, processing each identified portion of the first set of audio signals to generate a corresponding subword unit representation of the identified portion, and forming a representation of a</u> ~~processing data representing a the first set of audio signals to form a representation of the~~ spoken event of interest, wherein the ~~processing~~ <u>forming</u> includes ~~receiving an indication that a spoken event in a first set of audio signals is of interest to a user, identifying two or more instances of the spoken event of interest in the first set of audio signals, and representing each identified instance of the spoken event of interest in the representation of the spoken event of interest using at least one sequence of subword units~~ <u>combining the subword unit representations of the respective identified portions of the first set of audio signals</u>;

a data storage for receiving the representation of the spoken event of interest from the speech recognizer;

a word spotter configured to use the representation of the spoken event of interest to locate putative instances of the spoken event of interest in data representing ~~the~~ unknown speech in a second audio signal.

19.   (Previously Presented) The system of claim 18, wherein the word spotter is further configured to identify time locations of the second audio signal at which the spoken event of interest is likely to have occurred based on a comparison of the data representing the unknown speech with the representation of the spoken event of interest.